

# One century of hydrological monitoring in two small catchments with different forest coverage

Manfred Stähli · Alexandre Badoux ·  
Andreas Ludwig · Karl Steiner ·  
Massimiliano Zappa · Christoph Hegg

Received: 29 January 2010 / Accepted: 29 September 2010 / Published online: 12 November 2010  
© Springer Science+Business Media B.V. 2010

**Abstract** Long-term data on precipitation and runoff are essential to draw firm conclusions about the behavior and trends of hydrological catchments that may be influenced by land use and climate change. Here the longest continuous runoff records from small catchments ( $<1 \text{ km}^2$ ) in Switzerland (and possibly worldwide) are reported. The history of the hydrological monitoring in the Sperbel- and Rappengraben (Emmental) is summarized, and inherent uncertainties in the data arising from the operation of the gauges are described. The runoff stations operated safely for more than 90% of the summer months when most of the major flood events occurred. Nevertheless, the absolute values of peak runoff during the largest flood events are subject to considerable uncertainty. The observed differences in average, base, and peak runoff can only partly be attributed to the substantial differences in forest coverage. This treasure trove of data can be used in various ways, exemplified here with an analysis of the generalized extreme value distributions of the two catchments. These distributions, and hence flood return periods, have varied greatly in the

course of one century, influenced by the occurrence of single extreme events. The data will be made publicly available for the further analysis of the mechanisms governing the runoff behavior of small catchments, as well as for testing stochastic and deterministic models.

**Keywords** Runoff · Hydrology · Small catchments · Forest coverage · Extreme value distribution

## Background

The discharge at the outlet of a hydrological catchment is the integral response of the area to rainfall, snowmelt, evapotranspiration, and catchment-internal water fluxes and storages. By analyzing the stream runoff, it is possible to assess the characteristic sensitivity of the catchment to precipitation (Kirchner 2009), which depends on the catchment's geology, soil, vegetation, topography, and size. The response of stream runoff to precipitation of variable intensity has been a classical subject of investigation for generations of hydrologists. One of the main motivations has been to develop the ability to forecast peak runoff during extreme rain storms (e.g., Young 2002). Other studies have explored rainfall–runoff relationships as a basis for sediment and solute transport in torrents and rivers (e.g., Rickenmann and

---

M. Stähli (✉) · A. Badoux · A. Ludwig ·  
K. Steiner · M. Zappa · C. Hegg  
Mountain Hydrology and Torrents,  
Swiss Federal Research Institute WSL,  
Zürcherstrasse 111, 8903 Birmensdorf, Switzerland  
e-mail: manfred.staehli@wsl.ch

Koschni 2010), as well as for habitat conditions of fish, invertebrates, and other aquatic organisms (e.g., Jensen and Johnsen 1999). More recently, the issue of low flow as a consequence of drought has become another focus in runoff research (Hisdal et al. 2001).

Measuring runoff is one of the fundamental methods in hydrology, and it has a long tradition worldwide. Along the Nile, for example, flood-level gauges (so called Nilometers) have been run for thousands of years (Hurst and Phillips 1931). In Switzerland, the longest data record of stream runoff concerns its largest river, the Rhine (since 1808, Ghezzi 1926). But no continuous measurements of small catchments of just a few square kilometers are available from the nineteenth century. In 1900, however, the Swiss Federal Research Institute WSL installed two gauging stations in the Emmental region (central Switzerland) and started a long-term observation resulting in the longest continuous runoff data series of small catchments in Switzerland—and possibly worldwide.

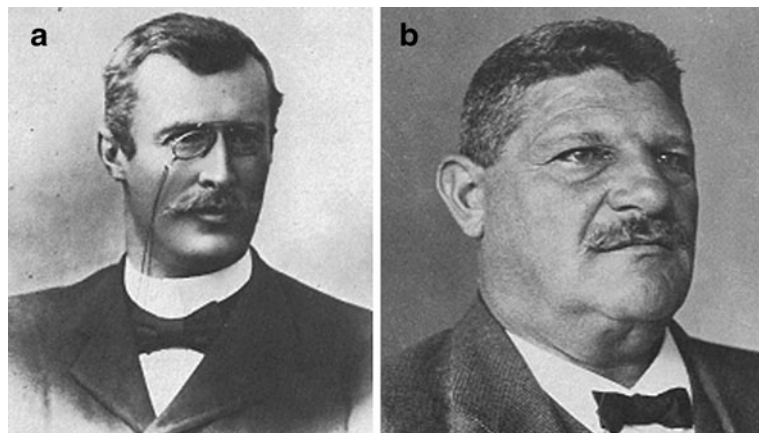
The motivation for recording these measurements had to do with the fact that Swiss forests had been in rather bad condition throughout the nineteenth century. A series of major flood events caused extensive damage in many areas (e.g., 1834, 1839, 1860, and 1868) and led to the hypothesis that over-harvested forests were one of the main reasons for the flooding (e.g., Landolt 1869). Several experts claimed that forests play a prominent role in mitigating such rainfall events. It was

hoped that studying the runoff behavior in the two catchments would throw light on this question, as one is completely forested (Sperbelgraben) and the other to a large extent covered by alpine grassland (Rappengraben).

The two study areas were proposed by the regional forest ranger, who had been involved in local afforestation projects. Professor C. Bourgeois (Fig. 1a), director of the EAFV (Eidgenössische Anstalt für das forstliche Versuchswesen, today WSL) at that time, promoted this idea and initiated the construction of the runoff stations. During the first decades, the measurements were operated by the local forest ranger under the supervision of the regional forest ranger and experts from EAFV. Local farmers were paid for minor jobs like opening channels in case of flooding.

The first results from these measurements were published by Engler (Fig. 1b; 1919) and Burger (e.g., 1934), who showed that the forested basin had lower peak flows, but higher baseflows. This confirmed the prevalent belief that forests helped to reduce flood flows and sustain baseflows. According to Engler (1919), the mitigating effect on peak discharge was especially high during short and intense thunderstorms, whereas only slight to no effects were observed during long-duration rainfall, depending on the water content of the soil before the event. This last finding has been referred to less often. Even though Engler stressed the fact that it was the forest soil rather than the vegetation cover that reduced flood peaks, this did not lead to changes in forest management. The

**Fig. 1** **a** Professor C. Bourgeois, director of the EAFV 1897–1901. **b** Professor A. Engler, director of the EAFV 1902–1923 (from Wullschleger 1985)



oversimplified result that “forests reduce floods” was what most influenced forest hydrology during the following decades when Burger was working on the data. Such detailed comparative analyses were abandoned in 1957, when the operation of the runoff and meteorological stations were taken over from the Swiss Federal Office for the Environment (FOEN) and the Federal Office of Meteorology and Climatology MeteoSwiss, respectively. These two institutions have ensured that the measurements have continued to the present day.

In North America, there has also been a lively debate about the influence of forest on the regulation of runoff in general and on flooding (Douglass and Hoover 1988). This research field has greatly benefited from the established forest management practice of large-scale clearcutting. The demand for quantitative data led the US Forest Services and the US Weather Bureau to start the first experiments measuring streamflow before and after tree removal at Wagon Wheel Gap (CO, USA) in 1909 (Bates and Henry 1928). Streamflow from two watersheds were monitored for 8 years, after which one watershed was cleared. The subsequent change in streamflow was attributed to the reduction in forest cover. In the following decades, many forest treatment experiments with controlled changes in forest cover were conducted throughout North America. Many of them are summarized in Hibbert (1967), and subsequently in reviews including international results by Bosch and Hewlett (1982) and Brown et al. (2005). Best known are probably the studies that have been carried out at Coweeta Hydrologic Laboratory (NC, USA) on about 25 small catchments since 1934 (Swank et al. 2001) and in the Hubbard Brook Experimental Forest (NH, USA; Hornbeck et al. 1997). Similar long-term data-sets of forest hydrology have also been recorded in other parts of the world, e.g., in Australia (e.g., Bren and Hopmans 2007) and New Zealand (e.g., Fahey and Jackson 1997).

One reservation with the Sperbel-/Rappengraben experiment is that the two catchments compared are not only forested differently, but might also differ in other ways, e.g., geomorphologically, in the topography of the soil–bedrock interface or in the width and

sequence of differently weathered bedrock layers. Even 50 years ago, Penman (1959) suspected some leakage from the Sperbelgraben, which was later supported in a modeling exercise (Badoux 2005) (cf. “Results and discussion”). In other forest hydrological experiments carried out mainly in the USA, such difficulties were avoided by carrying out control measurements for years in several catchments before changing the forest cover in some catchments and leaving others untouched. However, the statistical interpretation of datasets from such paired watershed experiments is complex and the inappropriate application of statistical methods has often produced erroneous conclusions in the past, as discussed in Alila et al. (2009).

This paper presents the complete runoff history of one century of the Sperbel-/Rappengraben catchments. Based on the valuable work of many other people, we have been able to compile a nearly complete digital data-set from 1903 to date, which may well be a unique source for future analyses. The inherent flaws of the original data-set have been, to the best of our knowledge, revised and the gaps “filled” with interpolation and simulated values. In spite of this careful rework, some uncertainty still remains, in particular with regard to peak discharge values.

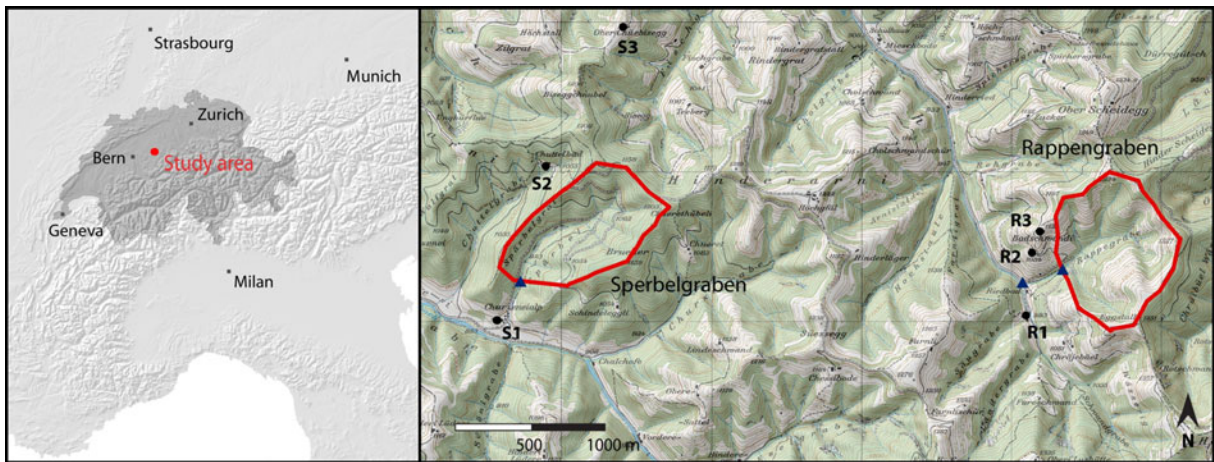
The aims of this paper are:

- to compile the history of the longest hydrological monitoring of small, partially forested catchments in Switzerland,
- to present the full data-set and point out the inherent uncertainties that need to be considered in subsequent analyses,
- to provide an example demonstrating the benefits of such a long data series, and
- to discuss any open research questions related to the runoff formation that these data might help to answer.

## Site and methods

### Description of the two catchments

The Sperbelgraben and Rappengraben catchments are situated in two parallel valleys of the



**Fig. 2** Map showing the location of the two catchments. The triangles denote the runoff stations (note that in 1927 the Rappengraben station was moved 300 m higher) and circles denote the rain gauges (S1 Kurzeneialp, S2 Kut-

telbad, S3 Bisegg, R1 Riedbad, R2 Lower Badschwendi, R3 Upper Badschwendi). Reproduced with permission of Swisstopo (JA 082265)

hilly Emmental region in the Swiss Pre-Alps (Fig. 2). The direct distance between the gauging stations is 3.6 km. The two catchments cover a comparable area of just over 0.5 km<sup>2</sup> (Table 1), and both drain from north-east to south-west. They have a similar channel density and an identical circumference, but differ in shape. The Rappengraben is quite circular while the Sperbelgraben is rather elongated. The main stream lengths amount to 1.01 km for the former and 0.76 km for the latter.

The catchments are situated in the molasse zone, a foreland basin north of the Alps, that formed during the Oligocene and Miocene epochs, and geologically consist of conglomer-

ate layers crossed by marl layers. While the clay content of the soils varies with the fraction of marl in the bedrock, their lime contents are generally low. The Sperbelgraben and Rappengraben are principally characterized by Cambisols, with an intermediate water-storage capacity and moderate permeability. Water-saturated soils, typically Gleysols, are largely restricted to the outcrop of the marl layers.

The Sperbelgraben catchment has an elevation range from 911 m a.s.l. at the Kurzeneialp gauging station to 1,203 m a.s.l. at its highest point, and is entirely covered with forest (Table 1). Main tree species include fir (*Abies alba*), spruce (*Picea abies*), beech (*Fagus sylvatica*), and sporadically

**Table 1** Catchment characteristics of the Sperbelgraben and Rappengraben, both located in Sumiswald, Canton Bern

		Sperbelgraben	Rappengraben
Area	[km <sup>2</sup> ]	0.544	0.596
Mean elevation	[m a.s.l.]	1,063	1,141
Highest elevation	[m a.s.l.]	1,203	1,256
Elevation gauging station	[m a.s.l.]	911	996
Circumference	[km]	2.96	2.97
Channel density	[km km <sup>-2</sup> ]	5.50	4.69
Fraction of forest cover	[%]	96.9 (1915) <sup>a</sup> 100 (2009) <sup>b</sup>	35.1 (1915) <sup>a</sup> 53.5 (2009) <sup>b</sup>
Fraction of alpine meadow and pasture	[%]	1.7 (1915) <sup>a</sup> 0.0 (2009) <sup>b</sup>	63.2 (1915) <sup>a</sup> 46.5 (2009) <sup>b</sup>
Dominant soil type		Cambisol	Cambisol

<sup>a</sup>Source: Burger (1934)

<sup>b</sup>Source: Federal Office of the Environment (<http://www.hydrodaten.admin.ch/e>)



maples (*Acer pseudoplatanus*). In contrast, forest covers only about half of the Rappengraben catchment (elevation range, 996 to 1,256 m a.s.l.). The rest of the catchment is used as alpine pasture. The fraction of forested areas, however, has been steadily increasing at the expense of the agricultural areas. In the first half of the twentieth century, approximately one third of the catchment was forested (Engler 1919; Burger 1934, 1943, 1954).

## Hydro-meteorological measurements

### *Runoff stations: period 1903–1927*

The beginning of measurement in the Sperbel- and Rappengraben is described in Engler (1919). The first runoff stations were built in 1900, but it was not until April 1903 that self-recording devices allowed continuous runoff observation in the two catchments. The so-called Bazin stations, as shown in Fig. 3a for the Rappengraben, were characterized by three parallel channels. Under low-flow conditions, only one of them was open. The second and third channels were activated manually with medium or high water discharge. Occasionally, the additional sections were not opened in time before a high flow event, and therefore the measurement was inaccurate. However, the affected measurements were subsequently corrected by applying a method outlined in Engler (1919, page 336). Although the calibrating curve was only determined for discharges lower than  $1.90 \text{ m}^3 \text{ s}^{-1}$ , the channel dimensions allowed the recording of discharge rates higher than  $3.50 \text{ m}^3 \text{ s}^{-1}$ .

### *Runoff stations: period 1927–1964*

In autumn 1927, new stations, each with a Thompson cross-section (an equal-leg triangle; Fig. 3b), were installed to replace the Bazin stations. Thus, operation became more convenient, as the setup was now independent of runoff. Moreover, the narrow Bazin channels often became clogged with wood and leaves (Burger 1943), which was another problem that was partly solved by the new cross-section. On the other hand, the maximum measurable discharge dropped to  $1.36 \text{ m}^3 \text{ s}^{-1}$ .

Building the new Rappengraben station presented the opportunity to shift its location 300 m upstream onto solid rock. The reason for this shift was to prevent a suspected loss of underground water at the original location in an area with sandy depositions next to the outlet into the river Hornbach (Burger 1943). The catchment's size thus decreased from 0.697 to 0.596 km<sup>2</sup>. At Sperbelgraben the runoff station remained at its original location.

### *Runoff stations: period 1964–2009*

Apparently, the two stations then underwent no relevant change until November 1964, when the Rappengraben runoff station was equipped with its present cross-section. It is characterized by a V-shaped geometry in the lowest 50 cm, which passes over into high, vertical walls on both sides. The capacity thus increased to  $2.14 \text{ m}^3 \text{ s}^{-1}$  (corresponding to a water level of 110 cm). Up until 1972, the calibration curve changed several times. While these variations were only minor in most cases, the depth–discharge relation changed peculiarly

**Fig. 3** **a** Bazin runoff station at Rappengraben 1903–1927 (from Engler 1919). **b** Present runoff station (Thompson cross-section) at Sperbelgraben (photo: Federal Office of the Environment)



in 1965. Although it could not be verified, this might indicate that the rebuilding was not finished until January 1966. Again the relationship changed noticeably in winter 1971/1972, in all likelihood due to the installation of a wooden invert at the bottom of the channel. The present calibration curve (from 2002) basically corresponds to that from 1972, with a maximum discharge of  $2.50 \text{ m}^3 \text{ s}^{-1}$  (corresponding to a water level of 120 cm). The most recent adjustment at the station was the installation of a tin invert from 1994.

The Thompson cross-section from 1927 is still at the Sperbelgraben station. In 1969, the capacity of  $1.36 \text{ m}^3 \text{ s}^{-1}$  was increased by  $0.18 \text{ m}^3 \text{ s}^{-1}$ . Alterations to the gauge well in September of the same year, however, unintentionally led to the floater becoming blocked as soon as a certain water depth was reached. This drawback was finally discovered and removed in October 1988. Therefore, no reliable discharge data  $>0.445 \text{ m}^3 \text{ s}^{-1}$  are available for that period. In August 2000, the 1927 cross-section was extended by 40-cm high, vertical walls on each side of the channel, resulting in an increased capacity of  $2.41 \text{ m}^3 \text{ s}^{-1}$  (Fig. 3b).

#### *Meteorological measurements*

Complementary to the runoff measurements, meteorological variables have also been recorded during the entire period, from 1903 to date. Until 1957, precipitation and snow depth were measured at three different locations in each catchment (indicated in Fig. 2). Readings of daily sums of precipitation were made at 7:30 a.m. using accumulative rain gauges. In addition, self-recording Hellmann precipitation gauges (Fig. 4) at the lowest stations of each catchment (i.e., Kurzeneialp and Riedbad) provided data with higher temporal resolution. Data gaps were scarce, and there was never more than one station with missing values at one time. During winter, snow depth was measured at a fixed measuring rod in close vicinity to each precipitation station. As for precipitation, the snow depth readings were taken at 7:30 a.m. It is important to note that the selected locations represented the snow cover in the open area without any forest influence. When measurements were first started, air temperature was measured



**Fig. 4** Rain gauge at Kurzeneialp, Sperbelgraben (from Engler 1919)

only at Kurzeneialp and Riedbad, and only later at all six sites. Three readings were taken each day in a shaded place with a thermometer produced by R. Fuess (Steglitz, Berlin).

After November 1957, all meteorological measurements were handed over to the Federal Office for Meteorology, the predecessor of MeteoSwiss. At this point, the meteorological measurements were reduced. For example, only the rain gauges of Kurzeneialp in the Sperbelgraben and Riedbad (1961–1973)/Lower Badschwendli (1974–present) in the Rappengraben were maintained. Snow depth measurements were completely abandoned, and air temperatures as well as other meteorological variables have been only available from nearby stations, such as the MeteoSwiss stations Napf and Langnau i.E.

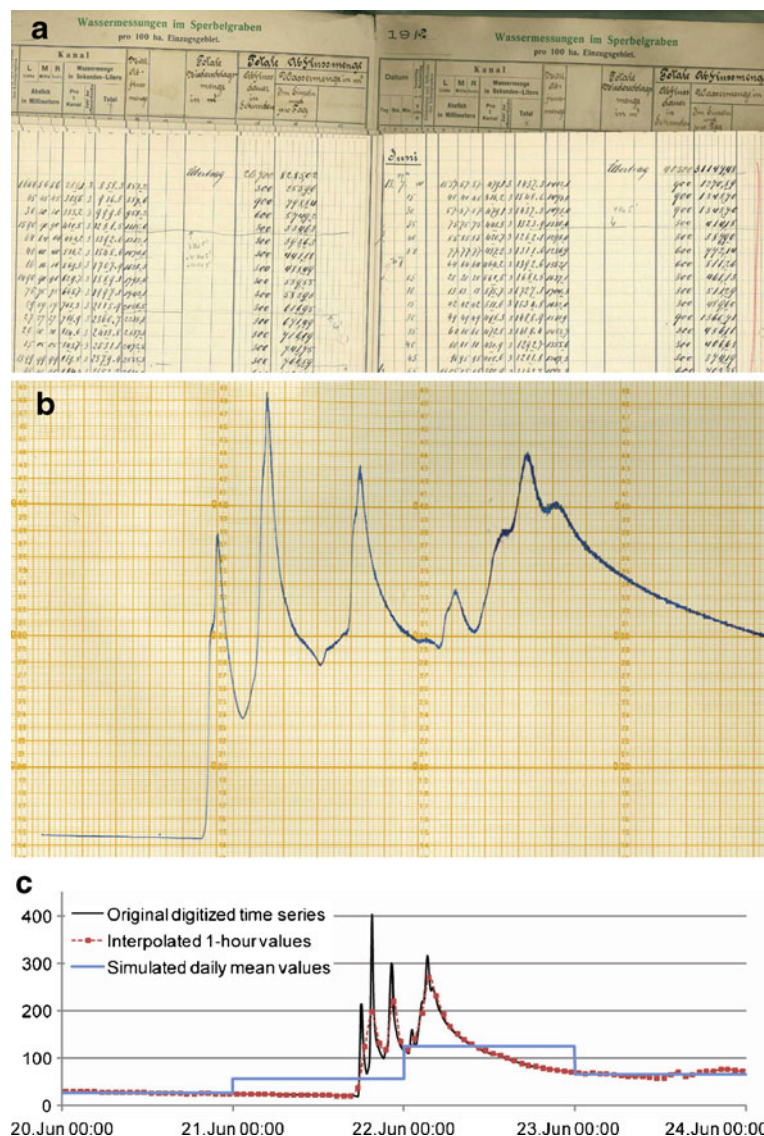
Unfortunately, the daily precipitation data from November 1957 to the end of 1960 in the two catchments has been lost. Instead, they have been reconstructed by temporally distributing monthly sums to daily values according to the occurrence and intensity of precipitation at three adjacent locations (Entlebuch, Affoltern i.E., and Lagnau i.E.).

# Runoff data-set and data processing

The original runoff data series for the period 1903 to 1957 were taken from historic handwritten books (Fig. 5a) at the Swiss Federal Institute WSL. The discharge values were transcribed manually from the hydrographs, which in turn were plotted automatically on measuring strips (Fig. 5b). The type of water-level recorder widely used in Swiss rivers and lakes at the time had been developed by the company Hasler AG (Engler 1919). Its basic components are a floater, a

marking device, and a clock. The temporal resolution of the transcription usually varies from 5 min during flood events up to 6 h at low discharge. Now, more than 50 years later, the runoff records from the books have been digitized. Since 1957, the Swiss Federal Office for the Environment has acquired the data. Until August 2001 in the Rappengraben and March 2003 in the Sperbelgraben, data were digitized manually based on the hydrographs. Since then, runoff data has been logged digitally. The logging systems were produced by the company Kern AG.

**Fig. 5** Examples of **a** original handwritten data; **b** a measuring strip recording the floating level; **c** digitized runoff curves at the original time resolution (black), interpolated to 1 h mean values (red dots), and simulated at a daily time resolution (blue)





Initially, a system called “Feldlogger FL-2” was installed. This was later replaced by the “Telelogger TL-1”. Analogue hydrographs are still plotted and used to validate the digital records. As in the first half of the twentieth century, the temporal resolution varies as a function of discharge, with time steps of less than a minute during flood events and several hours with stable runoff. For data applications that need a uniform time step (e.g., 10-min or 1-h values), we finally adopted linear interpolation or averaging if at least one daily measurement was available (Fig. 5c).

To obtain two complete 106-year long data series, we decided to fill the major gaps in the series by applying a state-of-the art semi-distributed rainfall–runoff model PREVAH (Precipitation–Runoff–Evapotranspiration–HRU) (Viviroli et al. 2009, see section below). In general, uncertain original data were not replaced with model values. However, where there were no data at all for more than 24 h, gaps were filled with simulated runoff values. Simulated runoff values were also used to correct the flawed measurements in the Sperbelgraben from 1970 to 1988 at discharge  $>0.445 \text{ m}^3 \text{ s}^{-1}$ , as discussed in the previous section.

To ensure comparability of the entire Rappengraben runoff series, the discharge measured before autumn 1927 was scaled to the catchment area of the gauging station built in 1927, which corresponds to the current official catchment area according to the Federal Office of the Environment and to the one indicated in Fig. 2 and in Table 1.

#### Completion of discharge time series by hydrological modeling

The hydrological model used for the completion of the runoff data-sets is the conceptually structured, semi-distributed model PREVAH. Its discretization relies on the aggregation of gridded spatial information into hydrological response units (HRUs; Gurtz et al. 1999). PREVAH consists of several different components, including a runoff generation sub-model and sub-models processing snow accumulation, snow melt, and glacier melt. A detailed model description including full information on the model

physics, structure, and parameterizations is given in Viviroli et al. (2009). Since the mid-1990s, PREVAH has been applied in various studies in Swiss catchments of variable size and in very different environments, both on a meso-scale (Gurtz et al. 2003; Zappa and Kan 2007) and a macro-scale (Zappa 2008). The present work, which is based on a previous simulation by Badoux (2005), is one of the first PREVAH applications on a lower meso-scale.

For the two catchments, the model has been forced by daily, spatially interpolated values of the following observed climatic variables: precipitation, air temperature, relative sunshine duration, wind speed, and relative humidity. The procedures adopted for spatial and temporal interpolation are based on elevation-dependent regressions and inverse distance weighting techniques (Viviroli et al. 2009). HRUs were extracted according to elevation zone, aspect, land use, and forest soil type (Badoux 2005).

Model calibration included the years 1975 to 1984, while 1974 was used as the initialization year. The calibration procedure relies on the monitored maximization of an acceptability score based on nine different objective functions (Zappa and Kan 2007). The functions test the agreement between observed and simulated discharges not only for the full calibration period, but also on a month-to-month and year-to-year basis (Viviroli et al. 2009). This allows a computationally effective identification of the parameter sets that provide a similar agreement in all the different seasons of the year and in all years of the calibration period. Subsequently, the Sperbelgraben catchment simulation was verified for the years 1904 to 1974 and 1985 to 2005. The Nash–Sutcliffe efficiency (e.g., Schaeffli and Gupta 2007) obtained for the Sperbelgraben is 0.63 for both calibration and verification periods. This is a rather fair value for such a small basin with little seasonality in discharge but with a high response of discharge to point rainfall. The Nash–Sutcliffe efficiency for the Rappengraben is poorer, but still indicates model skill. In the calibration period, the Nash–Sutcliffe efficiency is 0.57, and in the verification period 0.40. The main reason for the less skillful simulation in the Rappengraben is related to the snowmelt season. Finally, the



calibrated model was applied for the period 1904 to 2005.

## Results and discussion

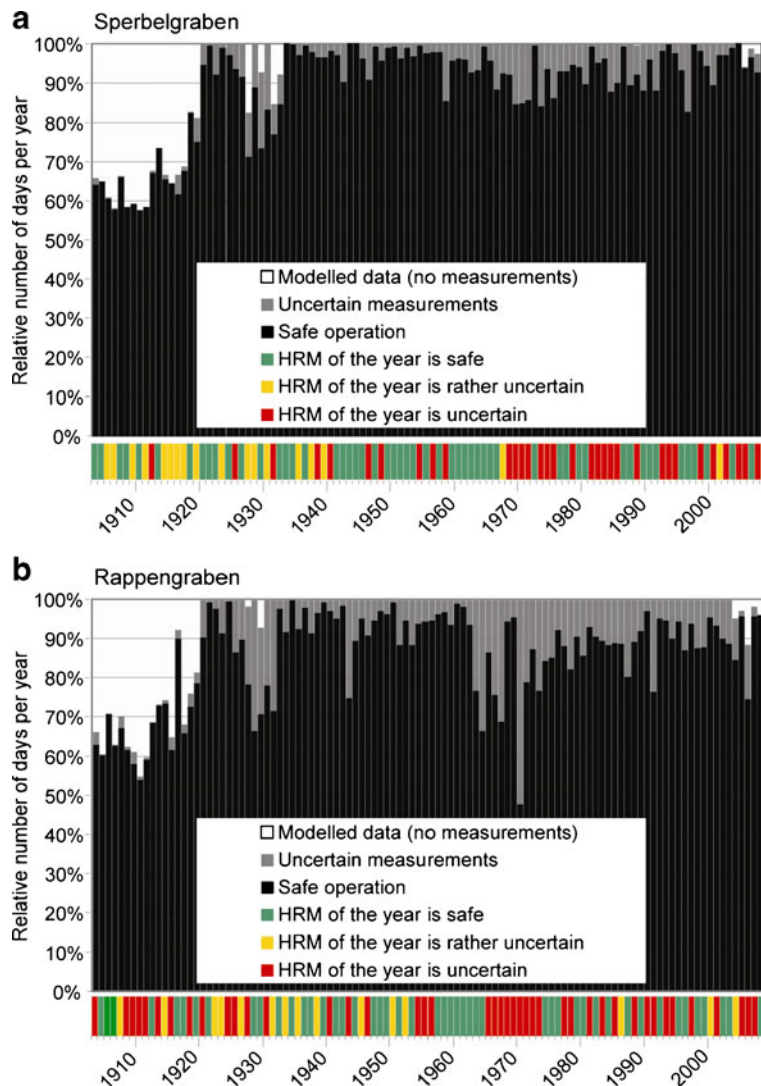
### Gaps and uncertainties in the runoff measurements

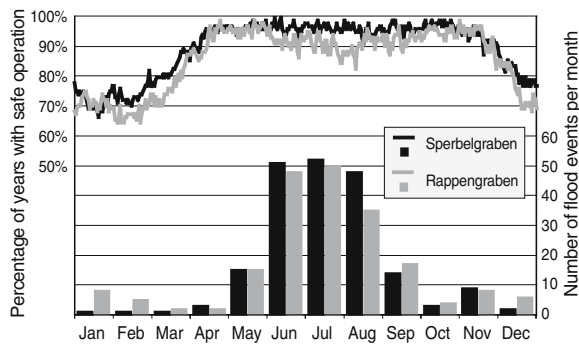
Clearly, the runoff data-set is not entirely homogenous over time due to changes in the station layout and recording methodology. In addition, there is uncertainty associated with a variety of problems that affected the measurements. During

flood events, insufficient capacity, bedload transport, and deposition have often led to inaccurate measurements. At lower discharge rates, uncertainty sporadically arises from natural factors, such as freezing in winter, or technical problems, e.g., clock malfunctioning. Original data during such periods is either inaccurate due to disturbed recording, interpolated (temporally, based on a regression with data from the second catchment), or just non-existent.

All the data books (1903–1957) and hydrographs (1957–2007) were verified for hints or comments on measurement problems. The results of this careful search are given in Figs. 6 and 7.

**Fig. 6** Overview of data availability over the entire measurement period 1903–2008 for **a** Sperbelgraben and **b** Rappengraben. The reliability of the highest runoff measurement (HRM) for each year is indicated in the colored bars below the main plots





**Fig. 7** Seasonal reliability of the runoff measurement system (*lines*) and seasonal occurrence of the 200 highest flood events (*bars*) during the measurement period 1903–2008

The former illustrates the percentage of days with flawless data, and those with uncertain or modeled data for each year. “Uncertain” here refers to runoff values that are indicated by corresponding comments in the data books or hydrographs. Figure 6 shows that prolonged gaps in data, where modeled data had to be used, are basically restricted to the period 1903–1919, when no measurements were taken during winter months (usually November/December to March/April). Additionally, the years 1929 (both catchments) and 1931 and 1932 (Sperbelgraben) were affected by the gauge freezing, whereas the gap in 1927 was due to the rebuilding of the station. Although winter runoff data have been collected in nearly all years since 1920, these were often based on a single daily measurement, resulting in a modest temporal resolution. In several other cases, including in recent decades, winter data missing in one catchment have been interpolated based on a regression with data from the other catchment, which leads to a higher proportion of uncertain data. For water balance studies and flood analyses, however, the reduced temporal resolution during winter is not critical since runoff is usually low and steady in winter, and flood events are rare (Fig. 7).

During summer months, the periods with uncertain data are shorter, but still occur from time to time. For the most part, summer-time data losses were caused by flood events. In the case of bedload transport and deposition, data collection

was impeded until the station could be cleaned or repaired, which varies from a single day to a couple of weeks. The years with uncertain maximum discharge values are indicated in Fig. 6.

Precise quantitative statements about the data quality cannot be made, as it is obviously impossible to determine the error of all the uncertain values in hindsight. However, the information in Figs. 6 and 7 allows us to make rough estimates. In winter, the percentage of years with uncertain or modeled data (Sperbelgraben 25–35%, Rappengraben 30–40%) is clearly higher than in summer (around 5% and 5–15%, respectively). Moreover, the figures imply that measurement problems have been more common in the Rappengraben than in the Sperbelgraben.

### Differences between catchments

In spite of the proximity of the two catchments (Fig. 2), the mean runoff in the Rappengraben is on average 32% greater than in the Sperbelgraben (period 1903–2005; Table 2). To some degree this can be explained by the differences in catchment size (Rappengraben 9.5% larger than Sperbelgraben; Table 1) and forest coverage, which leads to considerably higher transpiration in the Sperbelgraben and thus less water available for streamflow. Penman (1963) argued that “... it may be that the well-wooded Sperbelgraben has an evaporation

**Table 2** Summary statistics for the two catchments

	Sperbelgraben	Rappengraben
Mean runoff (1903–2005)	0.0155 m <sup>3</sup> s <sup>-1</sup>	0.0205 m <sup>3</sup> s <sup>-1</sup>
Coefficient of variation for daily runoff (1903–2005)	1.249	1.406
Base flow (Q <sub>347</sub> <sup>a</sup> )	0.00196 m <sup>3</sup> s <sup>-1</sup>	0.00180 m <sup>3</sup> s <sup>-1</sup>
Mean annual precipitation (1903–1999) <sup>b</sup>	1,636 mm	1,713 mm
Mean annual runoff (1903–1999) <sup>b</sup>	896 mm	1,084 mm

<sup>a</sup>Q<sub>347</sub> is a low-flow parameter corresponding to the runoff, which is exceeded 95% of the time

<sup>b</sup>Excluding 1957–1960, for which no daily precipitation data are available from the two catchments

loss some 10 to 15% greater than that of the dominantly grassland Rappengraben.” But the 4.6% more precipitation in the Rappengraben than in the Sperbelgraben (Table 2) also contributed to this difference in runoff. However, a simple water balance calculation suggests that an unknown amount of seepage into the ground, which is not captured by the runoff gauges, also takes place. The larger difference between precipitation and runoff in the Sperbelgraben ( $740 \text{ mm year}^{-1}$ ) compared with in the Rappengraben ( $629 \text{ mm year}^{-1}$ ) simply cannot be attributed to evapotranspiration alone. According to the Hydrological Atlas of Switzerland, HADES, a typical order of magnitude for the annual evapotranspiration in the Emmental amounts to 550–600 mm (Menzel et al. 1999). Badoux (2005) estimated the water loss due to seepage in the Sperbelgraben to amount to roughly  $200 \text{ mm year}^{-1}$ .

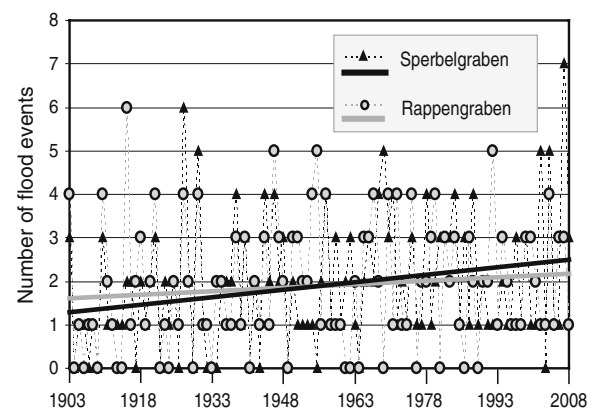
The coefficients of variation for daily runoff in the two catchments (Table 2) reflect the huge temporal variation, which is typical for such small, fast-responding catchments. They also indicate that the Rappengraben discharge is more dynamic than that of the Sperbelgraben, which—again—is partly the consequence of Rappengraben having considerably less forest.

As regards low flow, early on Engler (1919) noticed that the Sperbelgraben carries more water during drought periods than the Rappengraben. In fact, complete exsiccation, i.e., a situation without water flowing in the stream, has only been observed in the Rappengraben so far. A comparison of the  $Q_{347}$  values (a  $Q_{347}$  value is a widely used low-flow parameter in Switzerland, corresponding to the runoff that is exceeded during 95% of the time) confirms the slightly (8%) higher base flow in the Sperbelgraben (Table 2). While Engler (1919) attributed this phenomenon to the positive effect of the vegetation, Robinson et al. (2003) suggest that the deeper soil in the Sperbelgraben is responsible for this effect. However, it remains unclear why they believe the soils in the Sperbelgraben to be deeper. Our soil mapping did not provide any evidence of a systematic difference in soil depth between the two catchments, but it did confirm that the geology and the soil types are similar. The fact that the number of marl layers is

larger in the Rappengraben than in the Sperbelgraben may be relevant for the runoff formation.

Peak flow can be more than three orders of magnitude higher than base flow in the two catchments, which is typical for steep torrents. The highest observed discharge values in the Rappengraben exceed those in the Sperbelgraben by a factor of 1.5 to 2.5. This will be discussed in detail in the following section. Based on the 106-year data-set, we investigated whether a trend in the occurrence of flood events can be observed over the last century (Fig. 8). Such a trend could either reflect the impact of a change in the vegetation or in the climate. To this end, we arbitrarily selected the 200 largest runoff events in each catchment and determined their occurrence for each year. For the Sperbelgraben that included all floods with peak discharge  $>0.315 \text{ m}^3 \text{ s}^{-1}$  and for the Rappengraben  $>0.473 \text{ m}^3 \text{ s}^{-1}$ . (Note that this analysis is not affected by the problem caused by the blocked floater in the Sperbelgraben from 1970 to 1988 because the threshold value is below the critical level of  $0.445 \text{ m}^3 \text{ s}^{-1}$ .) Figure 8 indicates there has been a slight increase in the occurrence of flood events in both catchments, in particular in the Sperbelgraben. However, this apparent positive trend is not statistically significant at the 95% confidence level.

Related to flood events is the question whether the annual highest runoff peaks have occurred



**Fig. 8** Temporal distribution of the 200 highest flood events during the measurement period 1903–2008 (the two straight lines show the linear regression of the number of flood events per year)

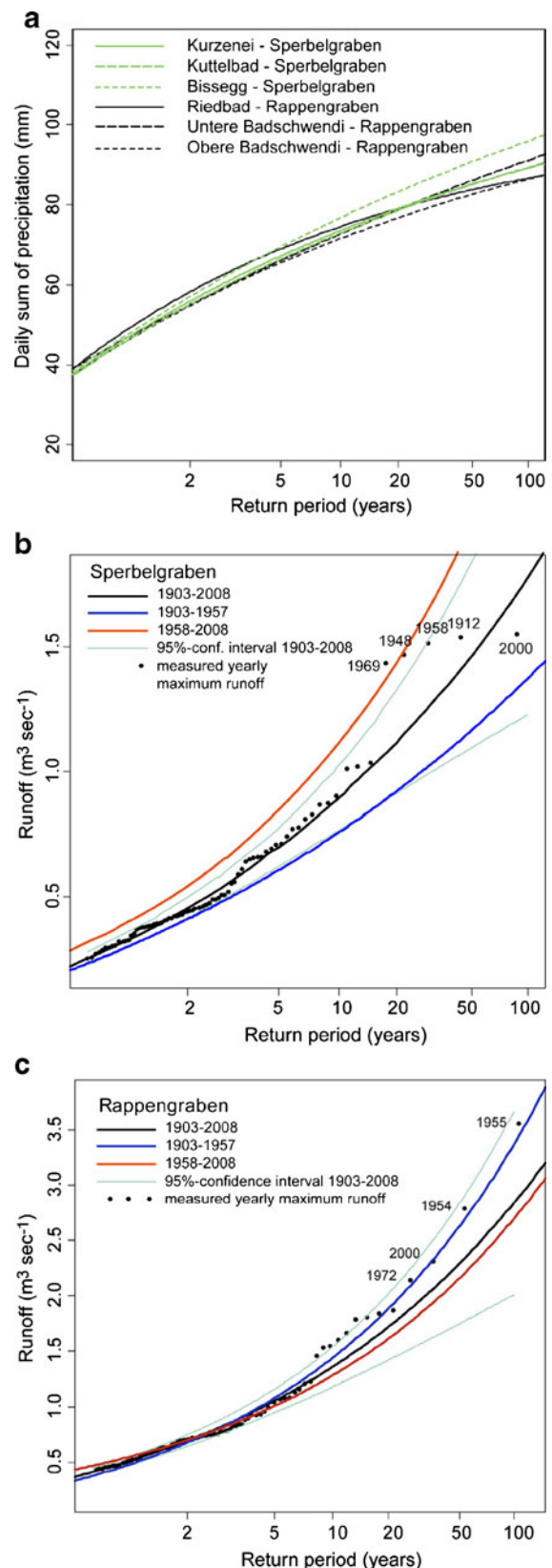
during the same storm event in both catchments or not. The present data-set suggests that only in 44% of the cases was the yearly maximum discharge triggered by the same rainfall event (i.e., the same day or 1 day delayed). This might seem surprising in view of the short spatial distance of roughly 3.5 km between the two catchments, but it indicates again that major thunderstorms in this area can emerge very locally and may strike small catchments at random. The very small spatial extent of thunderstorms and its significance for catchment runoff has been observed worldwide in several studies, e.g., in Goodrich et al. (1995).

Using the data-set to analyze the generalized extreme values distribution

The use of the data-set can be illustrated with an analysis of the generalized extreme values distribution, GEVD, for the runoff in the two catchments. This extreme value statistic is a widely used tool for practitioners to design flood protection constructions, culverts, and passages. Flood return periods (e.g.,  $Q_{100}$ , which in a statistical sense corresponds to a runoff peak occurring only once a century) are typically estimated from time series that are considerably shorter.

We plotted the GEVD for the runoff in the two catchments (Fig. 9) based on the full data-set from 1903 to 2008, as well as from sub-sets of approximately 50 years, representing the first and the second half of the century. We used a Gumbel distribution to fit the yearly maxima as proposed by Coles (2001).

The GEVD varies less in the Rappengraben than in the Sperbelgraben (Fig. 9c, b). This has most probably to do with the largest runoff measurements in the Sperbelgraben, some of which exceed the 95% confidence interval of the GEVD of the whole data-set. Notes in the historic records suggest that some of these values were overesti-



**Fig. 9** Generalized extreme value statistics for **a** daily sums of precipitation, **b** runoff in the Sperbelgraben, and **c** runoff in the Rappengraben. For precipitation, we used data from 1903–1956, and for runoff we used data from different periods (1903–1957: stations operated by WSL; 1958–2008: stations operated by FOEN)



mations of the true runoff due to the associated massive bedload transport.

Yearly maximum runoff in the Rappengraben exceeds that in the Sperbelgraben by a factor of 1.5 to 2.5. Since the GEVD of daily precipitation indicates no tendency toward systematically higher rainfall maxima in the Rappengraben (Fig. 9a), this difference in the GEVD for runoff can only be attributed to differences in the area's shape, forest coverage, geology, and the distribution of soil types. To what extent the suspected seepage loss contributed to this substantial difference in the GEVD in the two catchments remains an open question.

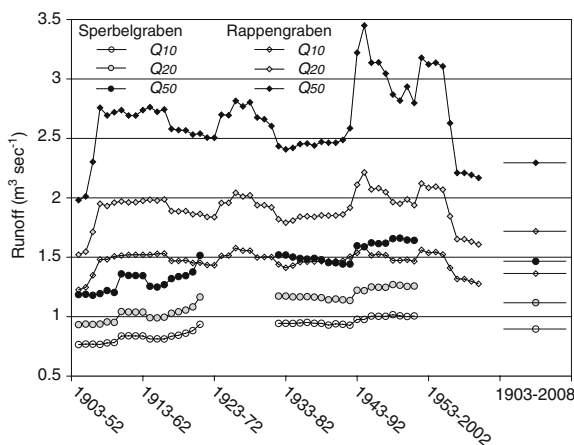
The present data-sets allow us to investigate the temporal variation of a peak flow with a given return period in the course of a century. We calculated  $Q_{10}$ ,  $Q_{20}$ , and  $Q_{50}$  from 50 consecutive yearly runoff maxima starting in 1903 and ending in 2008 for both catchments, assuming again a Gumbel distribution of the data. (In the Sperbelgraben the years between 1970 and 1988 had to be excluded due to the malfunction of the floater at runoff  $>0.445 \text{ m}^3 \text{ s}^{-1}$ . We therefore used the years before and after the gap to compose 50-year data-sets.) Figure 10 shows clearly how the peak flows with a given return period vary greatly in the course of a century. For example,  $Q_{10}$ ,  $Q_{20}$ , and  $Q_{50}$  increased considerably from the beginning to the second half of the twentieth cen-

tury. This is particularly conspicuous for  $Q_{50}$  in the Rappengraben. Apparently, a few single very large peak floods, such as the ones in 1954 and 1955, have a large impact on these statistics. Thus, the underrepresentation of large floods in the first decades of the twentieth century would have led to a severe underestimation of a 50- or 100-year flood later on. The flood frequency fluctuations observed in the two Emmental catchments support the results of Schmocker-Fackel and Naef (2010), who analyzed such variations on a national and a European scale, and concluded that they can be explained by changes in large-scale atmospheric circulation.

### Conclusions and open questions for future research

Acquiring a runoff record of this quality over such a time span in two small, highly dynamic catchments is thanks to the deserving care, perseverance, and vision of several generations of researchers. Looking back on the history of the Sperbel- and the Rappengraben, we can admire the way the responsible institutions and persons succeeded in keeping up a high level of performance/operation over more than 100 years. The data-set that we have today is continuous and homogeneous in spite of the impact of numerous floods and harsh weather conditions, and in spite of the people and institutions responsible changing.

However, some gaps and errors are inevitable in such a long record. The Sperbel- and Rappengraben stations managed, during this observation period, to operate without any problems more than 80% of the time. During summer months, when most of the large runoff events take place, the performance tends to be even higher ( $>90\%$ ). Nevertheless, some of the absolute values of peak runoff during the largest flood events may be subject to uncertainty due to the associated bedload transport and the undefined measurement cross-section at very high water levels. As shown above, this may have a considerable impact on the generalized extreme value distribution, which typically provides the basis for flood engineering.



**Fig. 10** Temporal variation in floods with a return period of 10 years ( $Q_{10}$ ), 20 years ( $Q_{20}$ ), and 50 years ( $Q_{50}$ ) from yearly maximum data of 50 consequent years

For the future, which research directions can be pursued with the available data-set?

In spite of the comprehensive and long-term work of the past century by Engler, Burger, Penman, Badoux, and others, there are still unsolved questions related to the different runoff behavior of the two catchments. However, the future focus may not be primarily on the difference in forest coverage, but rather on differences in their underground and soil type compositions. It appears that seepage water loss, which was already suspected by Penman (1959) and later supported by the numerical modeling of Badoux (2005), is not only a key factor affecting the water balance of the two catchments, but also their proneness to flooding. While Badoux (2005) estimated an average water loss of approximately 200 mm year<sup>-1</sup> in the Sperbelgraben, it is still unknown whether there is a similar seepage loss in the Rappengraben. Further, it would be essential to find out whether this water loss occurs continuously, or if it is restricted to rain and snowmelt situations. We presume that the spatial distribution and connectivity of the marl layers in the two catchments is decisive for the water loss and the fast runoff response to rainfall. A detailed geological mapping may illuminate this hypothesis.

Another use of the data may be in the context of conceptualization of runoff formation. It has been claimed by leading hydrologists that progress in the general understanding of the main controls of runoff formation can only be achieved through a more systematic and consistent analysis and comparison of different catchments (Tetzlaff et al. 2008). The present data-set, with its extensive spatial information on forest coverage, soil type distribution (available for Sperbelgraben), and high-resolution digital terrain information, as well as with small-scale soil hydrological data (Badoux et al. 2006), could be a valuable source for testing general concepts. Completing the existing data with tracer experiments would be helpful in determining the typical catchment descriptors of hydrological functioning, such as mean residence time (Soulsby and Tetzlaff 2008) or catchment transit times (Tetzlaff et al. 2009).

Further, these data allow evaluating the applicability of hydrological models over periods

longer than a few decades. This will probably raise needs for improved parameterizations of a model's land surface scheme. Such parameterizations would be important to assess changes in hydrological regimes under climate change. The presented data-set of the Sperbel- and Rappengraben perfectly suits for evaluating new models accounting for transient changes in forest coverage.

Finally, the 106-year Sperbel- and Rappengraben data will also offer opportunities to assess hydrological trends in such small catchments attributed to changes in forest coverage or meteorological conditions. It will be interesting to check the recent data to see whether the sensitivity of the runoff to rainfall (Kirchner 2009) has changed in the Rappengraben over the last century as a consequence of the forest expanding from 35% at the beginning of the nineteenth century to 53% at the beginning of the twentieth century (Table 1). In the context of climate change, the question of whether low flow or drought situations have become more frequent will also be interesting to explore.

Such exploration is encouraged and for this purpose the Sperbel- and Rappengraben data have been made publicly available at [www.wsl.ch/data/sperbel-rappengraben](http://www.wsl.ch/data/sperbel-rappengraben).

**Acknowledgements** The authors would like to thank the numerous known and unknown people who have contributed over the past century to this comprehensive data-set. In particular, we are grateful to the staff of the Swiss Federal Office of the Environment, MeteoSwiss, and the Swiss Federal Research Institute WSL and their predecessor institutions. Thanks are also due to Luzi Bernhard for support with data handling and Silvia Dingwall for language correction. The comments of three anonymous reviewers helped substantially improving the manuscript.

## References

- Alila, Y., Kuràs, P. K., Schnorbus, M., & Hudson, R. (2009). Forest and floods: A new paradigm sheds light on age-old controversies. *Water Resources Research*, 45, W08416. doi:10.1029/2008WR007207.
- Badoux, A. (2005). *The influence of wind storm deforestation on the runoff generation at various scales in a torrential catchment* (p. 125). PhD thesis, University of Bern. [http://www.stub.unibe.ch/download/eldiss/05badoux\\_a.pdf](http://www.stub.unibe.ch/download/eldiss/05badoux_a.pdf).

- Badoux, A., Witzig, J., Germann, P. F., Kienholz, H., Lüscher, P., Weingartner, R., & Hegg, C. (2006). Investigations on the runoff generation at the profile and plot scales, Swiss Emmental. *Hydrological Processes*, 20(2), 377–394.
- Bates, C. G., & Henry, A. J. (1928). Second phase of streamflow experiment at Wagon Wheel Gap, Colorado. *Monthly Weather Review*, 56(3), 79–85.
- Bosch, J. M., & Hewlett, J. D. (1982). A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration. *Journal of Hydrology*, 55, 3–23.
- Bren, L., & Hopmans, P. (2007). Paired catchments observations on the water yield of mature eucalypt and immature radiata pine plantations in Victoria, Australia. *Journal of Hydrology*, 336, 416–429.
- Brown, A. E., Zhang, L., McMahon, T. A., Western, A. W., & Vertessy, R. A. (2005). A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation. *Journal of Hydrology*, 310(1–4), 28–61.
- Burger, H. (1934). Einfluss des Waldes auf den Stand der Gewässer; 2. Mitteilung; Der Wasserhaushalt im Sperbel- und Rappengraben von 1915/16 bis 1926/27. *Mitteilungen der Schweizerischen Anstalt für das forstliche Versuchswesen*, 18(2), 311–416.
- Burger, H. (1943). Einfluss des Waldes auf den Stand der Gewässer; 3. Mitteilung; Der Wasserhaushalt im Sperbel- und Rappengraben von 1927/28 bis 1941/42. *Mitteilungen der Schweizerischen Anstalt für das forstliche Versuchswesen*, 23(1), 167–222.
- Burger, H. (1954). Einfluss des Waldes auf den Stand der Gewässer; 5. Mitteilung; Der Wasserhaushalt im Sperbel- und Rappengraben von 1942/43 bis 1951/52. *Mitteilungen der Schweizerischen Anstalt für das forstliche Versuchswesen*, 31(1), 9–58.
- Coles, S. (2001). *An introduction to the statistical modelling of extreme values*. London: Springer.
- Douglass, J. E., & Hoover, M. D. (1988). History of Coweeta. In W. T. Swank, & D. A. Crossley Jr. (Eds.), *Forest hydrology and ecology at Coweeta* (pp. 17–31). New York: Springer.
- Engler, A. (1919). Einfluss des Waldes auf den Stand der Gewässer. *Mitteilungen der Schweizerischen Anstalt für das forstliche Versuchswesen*, 12, 1–626.
- Fahey, B. D., & Jackson, R. (1997). Hydrological impacts of converting native forests and grasslands to pine plantations, South Island, New Zealand. *Agricultural and Forest Meteorology*, 84, 69–82.
- Ghezzi, C. (1926). Die Abflussverhältnisse des Rheins in Basel. *Mitteilungen des Amtes für Wasserwirtschaft* (Vol. 19). Bern.
- Goodrich, D. C., Faurés, J.-M., Woolhiser, D. A., Lane, L. J., & Sorooshian, S. (1995). Measurement and analysis of small-scale convective storm rainfall variability. *Journal of Hydrology*, 173, 283–308.
- Gurtz, J., Baltensweiler, A., & Lang, H. (1999). Spatially distributed hydrotopo-based modelling of evapotranspiration and runoff in mountainous basins. *Hydrological Processes*, 13(17), 2751–2768.
- Gurtz, J., Zappa, M., Jasper, K., Lang, H., Verbunt, M., Badoux, A., & Vitvar, T. (2003). A comparative study in modelling runoff and its components in two mountainous catchments. *Hydrological Processes*, 17(2), 297–311.
- Hibbert, A. R. (1967). Forest treatment effects on water yield. In W. E. Sopper, & H. W. Lull (Eds.), *Forest Hydrology* (pp. 527–543). Oxford: Pergamon.
- Hisdal, H., Stahl, K., Tallaksen, L. M., & Demuth, S. (2001). Have streamflow droughts in Europe become more severe or frequent? *International Journal of Climatology*, 21(3), 317–333.
- Hornbeck, J. W., Martin, C. W., & Eagar, C. (1997). Summary of water yield experiments at Hubbard Brook Experiment Forest, New Hampshire. *Canadian Journal of Forest Research*, 27, 2043–2052.
- Hurst, H. E., & Phillips, P. (1931). *General description of the basin, meteorology, topography of the White Nile basin. The Nile basin* (Vol. I). Cairo: Government Press.
- Jensen, A. J., & Johnsen, B. O. (1999). The functional relationship between peak spring floods and survival and growth of juvenile Atlantic Salmon (*Salmo salar*) and Brown Trout (*Salmo trutta*). *Functional Ecology*, 13(6), 778–785.
- Kirchner, J. W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research*, 45, W02429. doi:10.1029/2008WR006912.
- Landolt, E. (1869). Die Wasserverheerungen in der Schweiz im September und Oktober 1886. *Schweizerische Zeitschrift für Forstwesen*, 20(1), 1–9.
- Menzel, L., Lang, H., & Rohmann, M. (1999). Mean annual actual evaporation. In *Hydrological atlas of Switzerland* (chapter 4.1). Bern, Switzerland: Swiss Federal Office for the Environment FOEN.
- Penman, H. L. (1959). Notes on the water balance of the Sperbelgraben and Rappengraben. *Mitteilungen der Schweizerischen Anstalt für das forstliche Versuchswesen*, 35(1), 99–109.
- Penman, H. L. (1963). *Vegetation and hydrology*. Farnham Royal: Commonwealth Agricultural Bureau.
- Rickenmann, D., & Koschni, A. (2010). Sediment loads due to fluvial transport and debris flows during the 2005 flood events in Switzerland. *Hydrological Processes*, 24, 993–1007.
- Robinson, M., Cognard-Plancq, A.-L., Cosandey, C., David, J., Durand, P., Führer, H.-W., et al. (2003). Studies of the impact of forests on peak flows and baseflows: A European perspective. *Forest Ecology and Management*, 186, 85–97.
- Schaeffli, B., & Gupta, H. V. (2007). Do Nash values have value? *Hydrological Processes*, 21(15), 2075–2080.
- Schmocker-Fackel, P., & Naef, F. (2010). Changes in flood frequencies in Switzerland since 1500. *Hydrology and Earth System Sciences*, 14(8), 1581–1594.
- Soulsby, C., & Tetzlaff, D. (2008). Towards simple approaches for mean residence time estimation in ungauged basins using tracers and soil distributions. *Journal of Hydrology*, 363, 60–74.

- Swank, W. T., Vose, J. M., & Elliott, K. J. (2001). Long-term hydrologic and water quality responses following commercial clearcutting of mixed hardwoods on a southern Appalachian catchment. *Forest Ecology and Management*, 143(1–3), 163–178.
- Tetzlaff, D., McDonnell, J. J., Uhlenbrook, S., McGuire, K. J., Bogaart, P. W., Naef, F., et al. (2008). Conceptualizing catchment processes: Simply too complex? *Hydrological Processes*, 22, 1727–1730.
- Tetzlaff, D., Seibert, J., McGuire, K. J., Laudon, H., Burns, D. A., Dunn, S. M., et al. (2009). How does landscape structure influence catchment transit time across different geomorphic provinces? *Hydrological Processes*, 23, 945–953.
- Viviroli, D., Zappa, M., Gurtz, J., & Weingartner, R. (2009). An introduction to the hydrological modelling system PREVAH and its pre- and post-processing-tools. *Environmental Modelling & Software*, 24(10), 1209–1222.
- Wullschleger, E. (1985). 100 Jahre Eidgenössische Anstalt für das forstliche Versuchswesen 1885–1985. Teil 1: Die Geschichte der EAFV. *Mitteilungen der Eidgenössischen Anstalt für das forstliche Versuchswesen*, 61, 1–630.
- Young, P. C. (2002). Advances in real-time flood forecasting. *Philosophical Transactions of the Royal Society London*, 360, 1433–1450.
- Zappa, M. (2008). Objective quantitative spatial verification of distributed snow cover simulations—An experiment for the whole of Switzerland. *Hydrological Sciences Journal*, 53(1), 179–191.
- Zappa, M., & Kan, C. (2007). Extreme heat and runoff extremes in the Swiss Alps. *Natural Hazards and Earth System Sciences*, 7, 375–389.